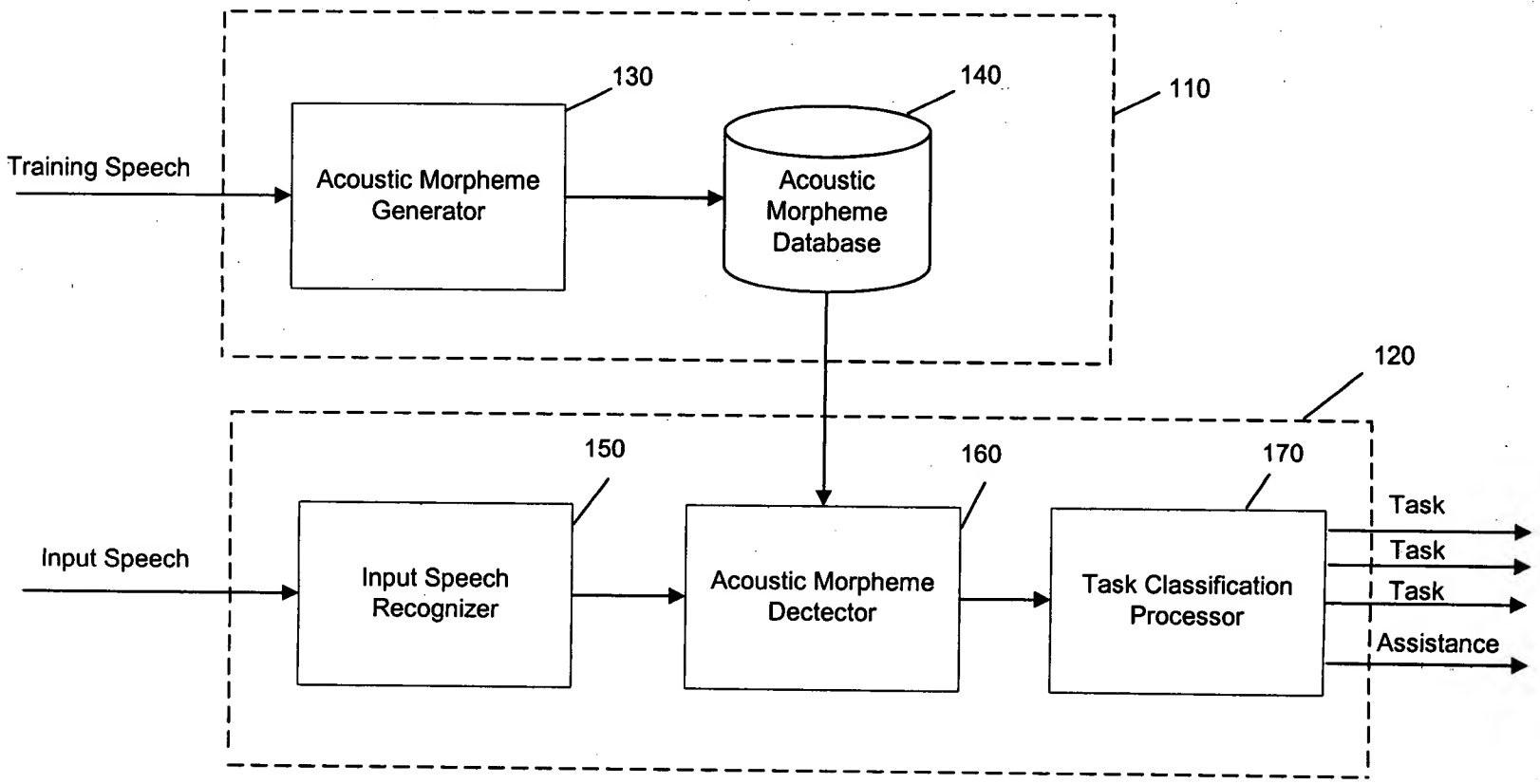

Fig. 1

09600721 _ 101900

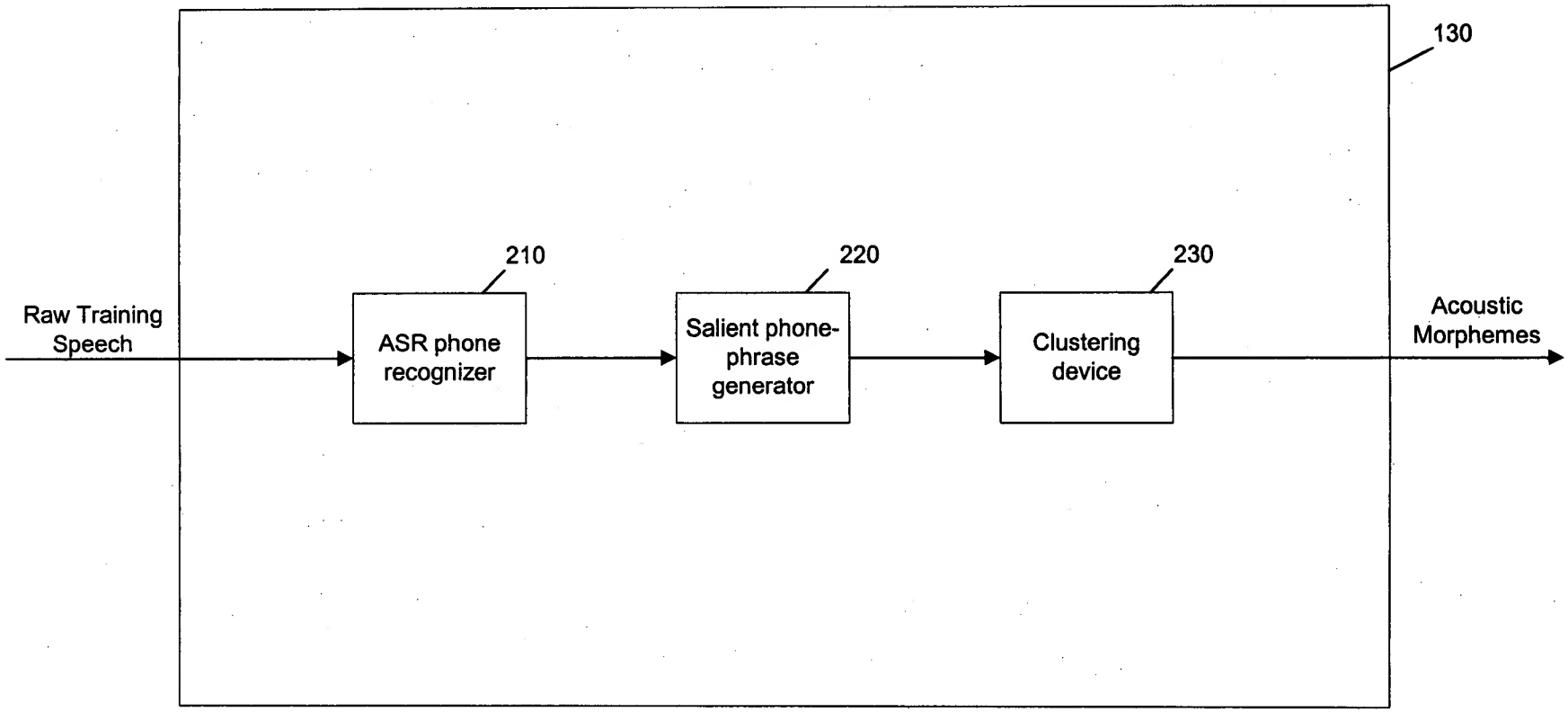

Fig. 2

0050072 1 . 1012000

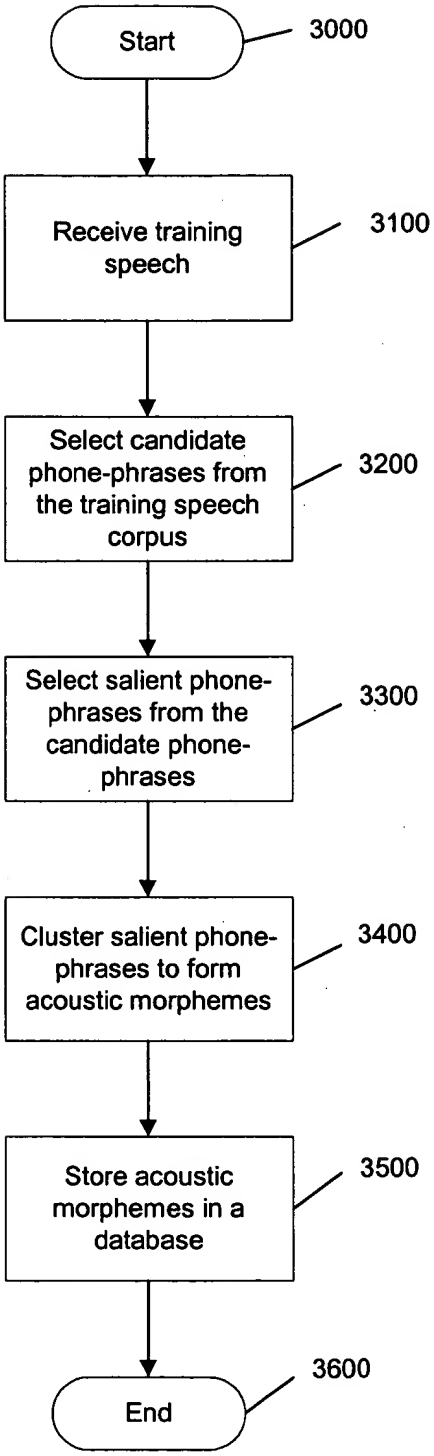
Start
3000
Receive training speech
3100
Select candidate phone-phrases from the training speech corpus
3200
Select salient phone-phrases from the candidate phone-phrases
3300
Cluster salient phone-phrases to form acoustic morphemes
3400
Store acoustic morphemes in a database
3500
End
3600

Fig. 3

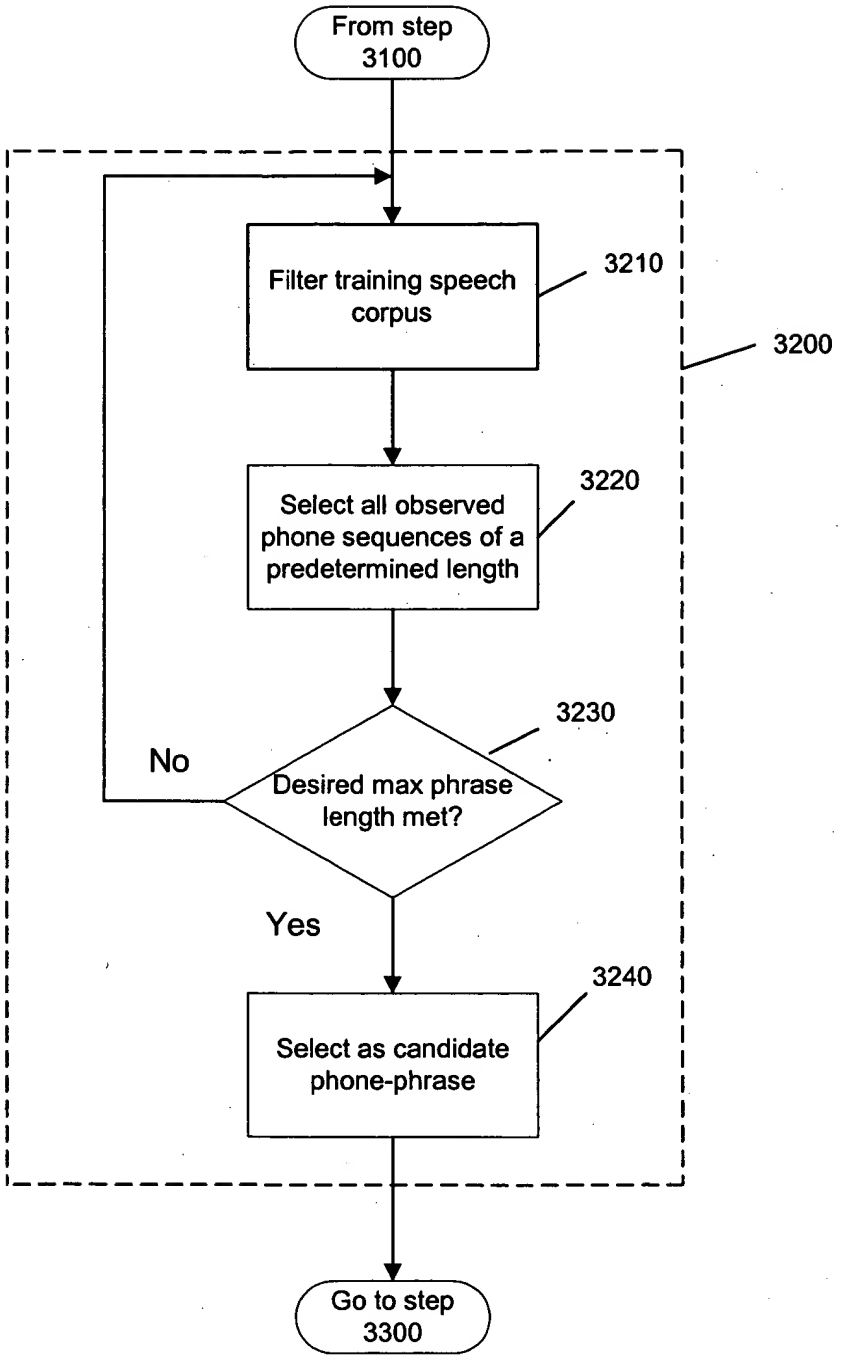
From step 3100
Filter training speech corpus
3210
3200
Select all observed phone sequences of a predetermined length
3220
3230
No
Desired max phrase length met?
Yes
3240
Select as candidate phone-phrase
Go to step 3300

Fig. 4

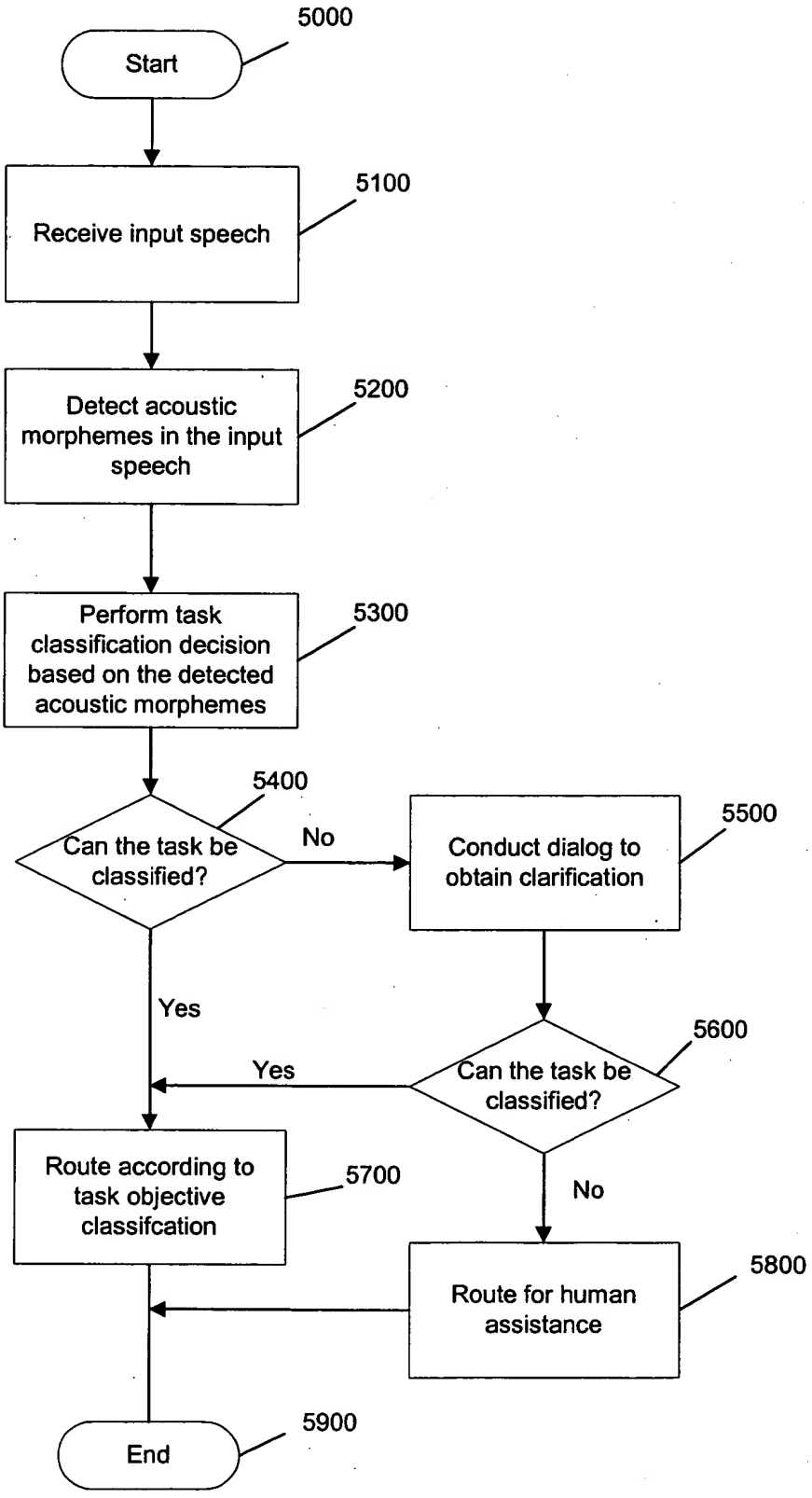
5000
Start
5100
Receive input speech
5200
Detect acoustic morphemes in the input speech
5300
Perform task classification decision based on the detected acoustic morphemes
5400
Can the task be classified?
No
5500
Conduct dialog to obtain clarification
Yes
5600
Can the task be classified?
Yes
Route according to task objective classifcation
5700
No
5800
Route for human assistance
5900
End

Fig. 5

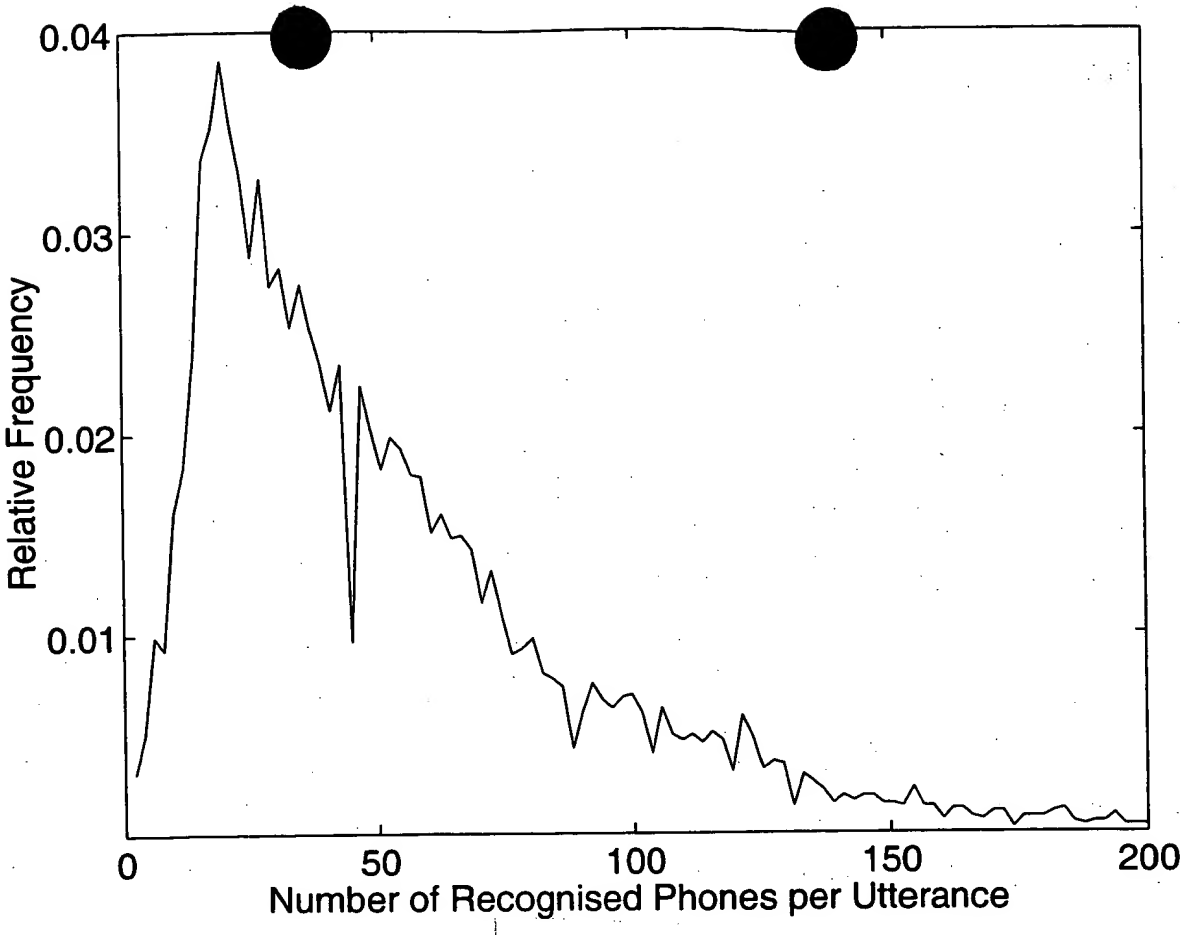
0.04
0.03
0.02
0.01
Relative Frequency
0
50
100
150
200
Number of Recognised Phones per Utterance

Fig. 6

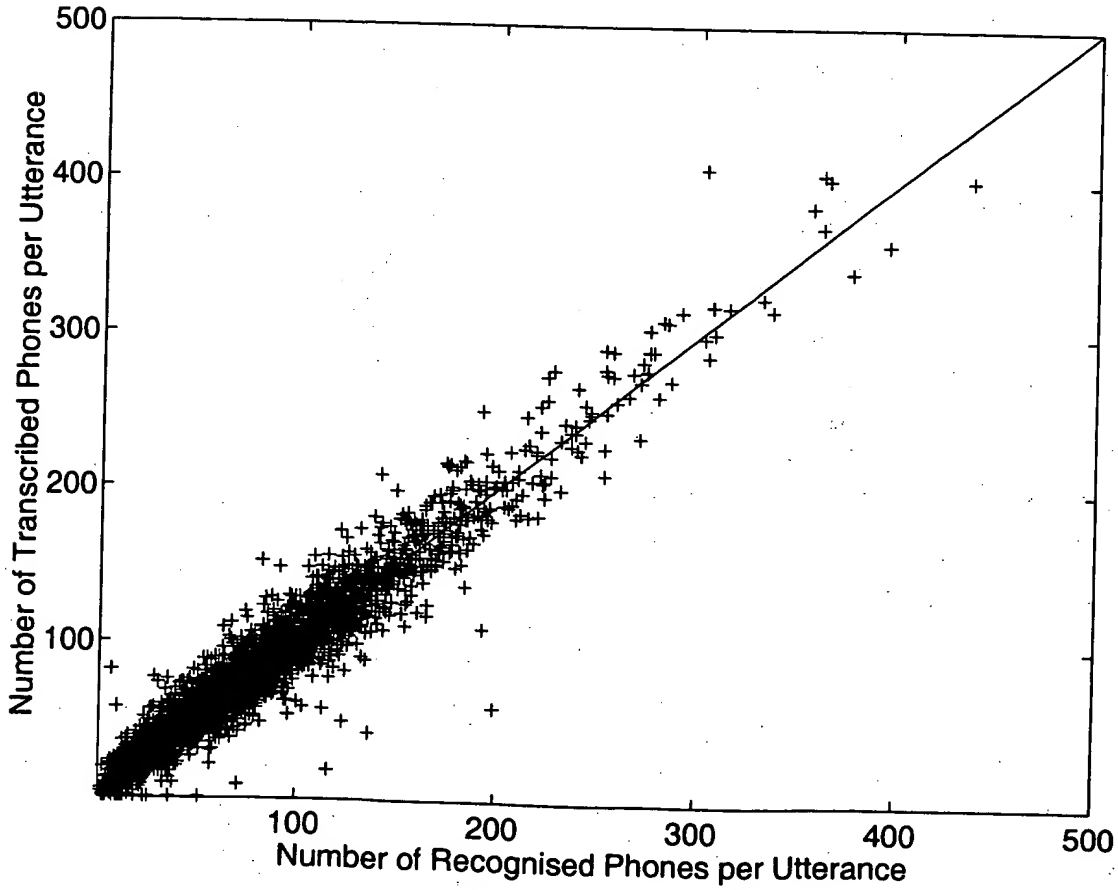
500
400
300
200
100
Number of Transcribed Phones per Utterance
100
200
300
400
500
Number of Recognised Phones per Utterance

Fig. 7

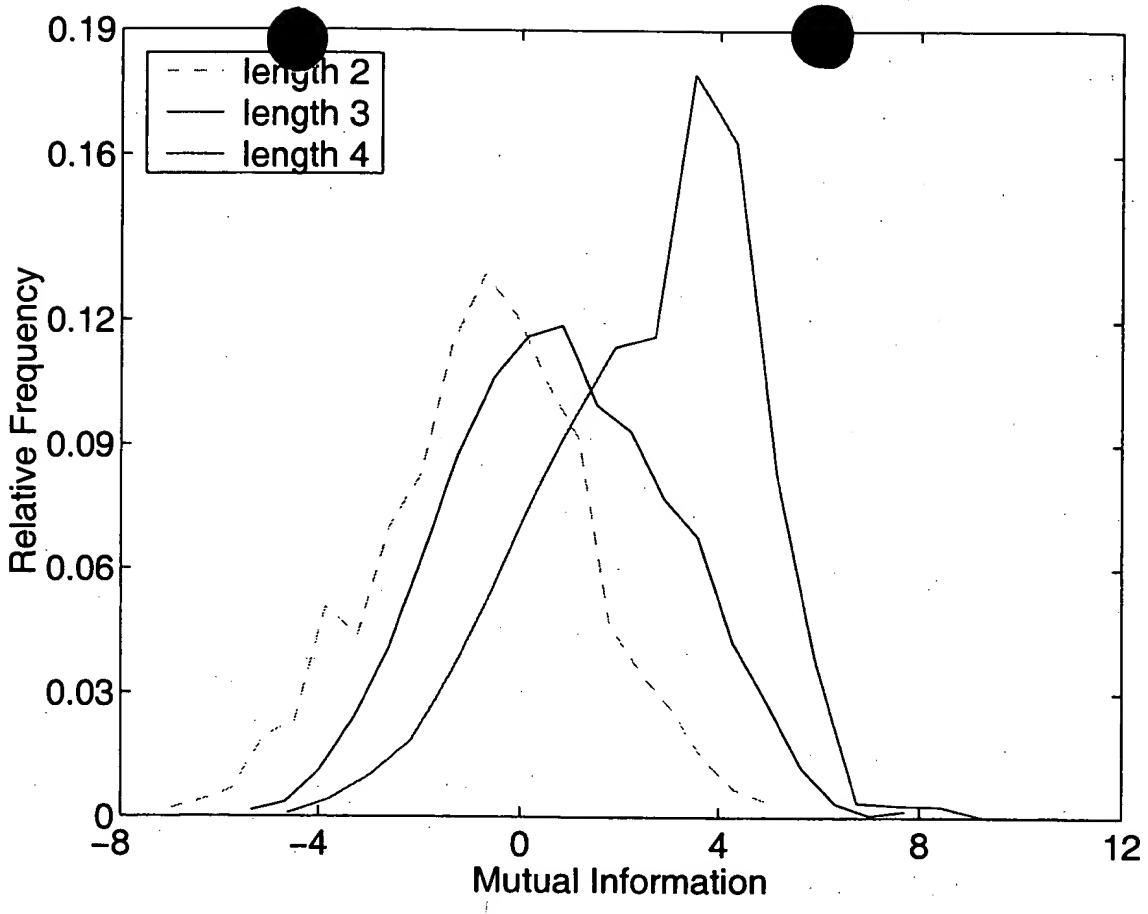
length 2
length 3
length 4
Relative Frequency
0.19
0.16
0.12
0.09
0.06
0.03
0
−8
−4
0
4
8
12
Mutual Information

Fig. 8

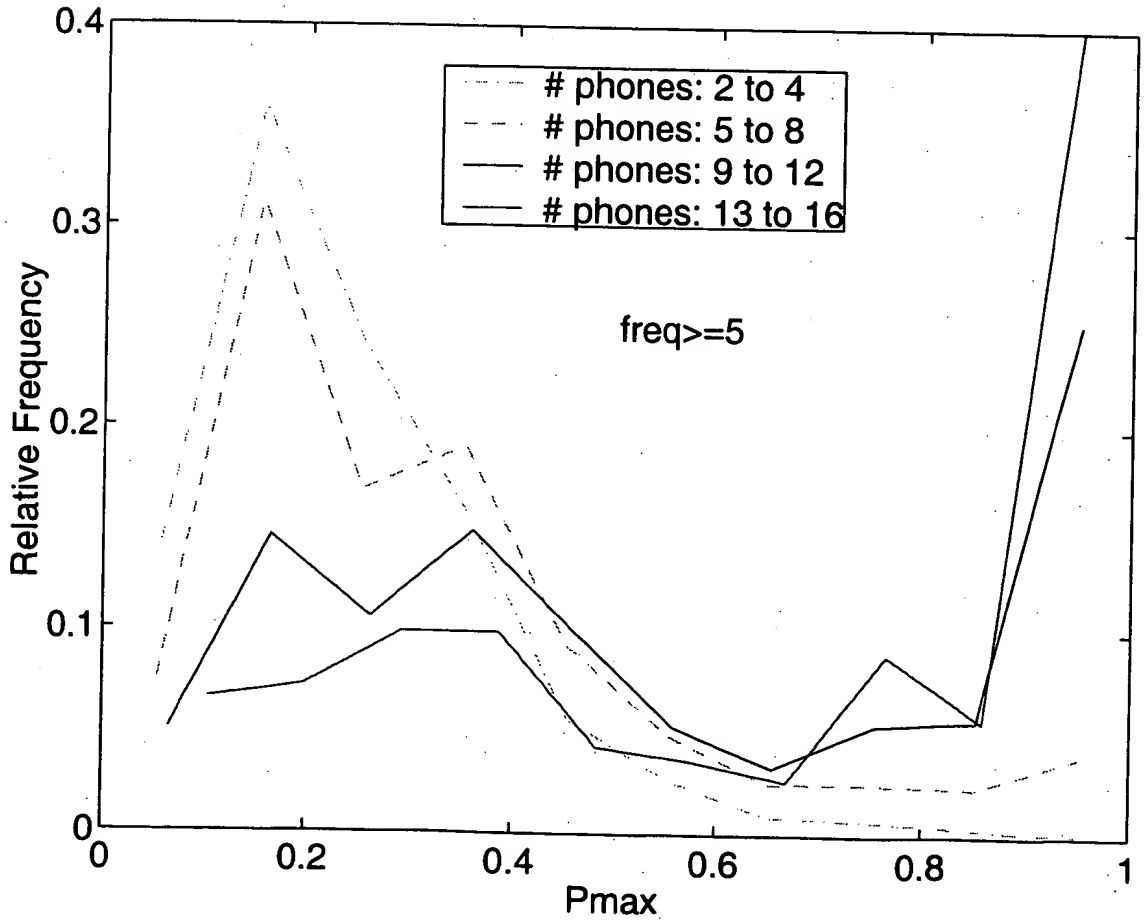
phones: 2 to 4
phones: 5 to 8
phones: 9 to 12
phones: 13 to 16
freq>=5
Relative Frequency
0.4
0.3
0.2
0.1
0
0
0.2
0.4
0.6
0.8
1
Pmax

Fig. 9

Pmax	Salient Phone Phrase
0.84	*K ax l eh K T*
0.95	*K ax l eh K T* K ao l
0.97	*K ax l eh K T* K ao l eos
0.98	m ey K ih T *K ax l eh K T*
0.94	*K l ay K* K ao l
0.93	*K l ay K* K ow l
0.90	P l ey s ih *K l ay K*
0.95	ih T *K ao r T* K ao l

Fig. 10

0068707 1E2

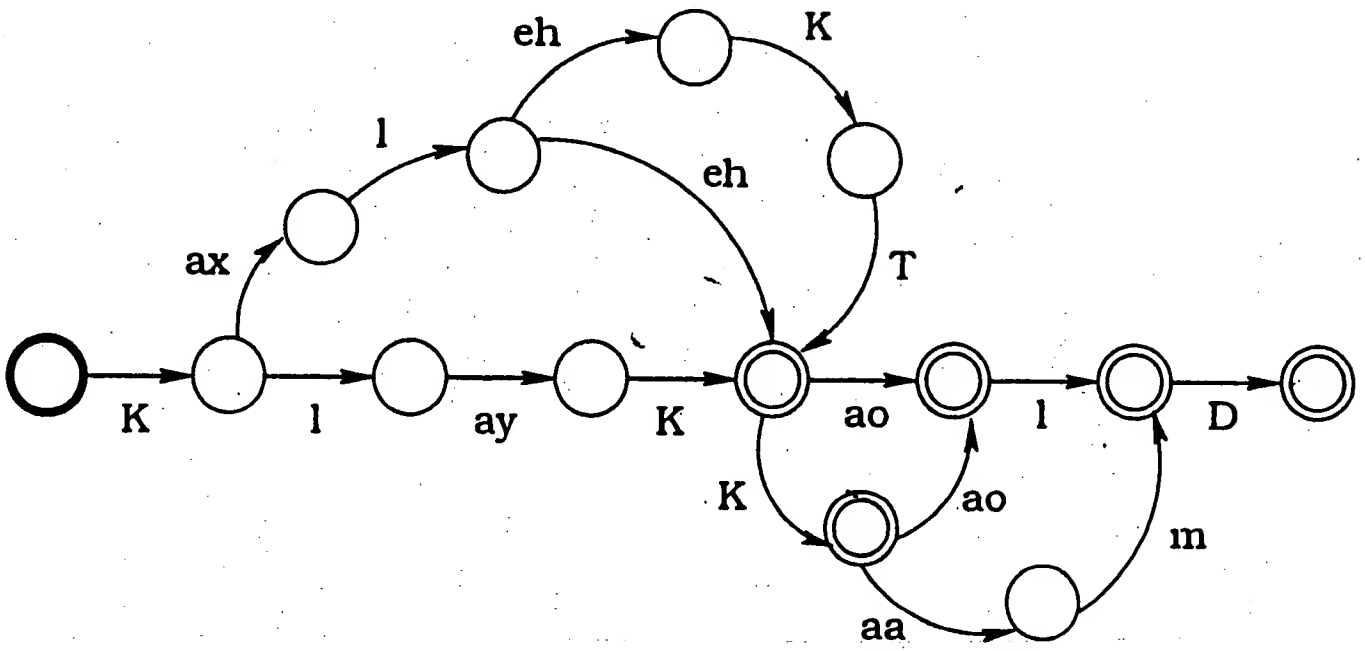

Fig. 11

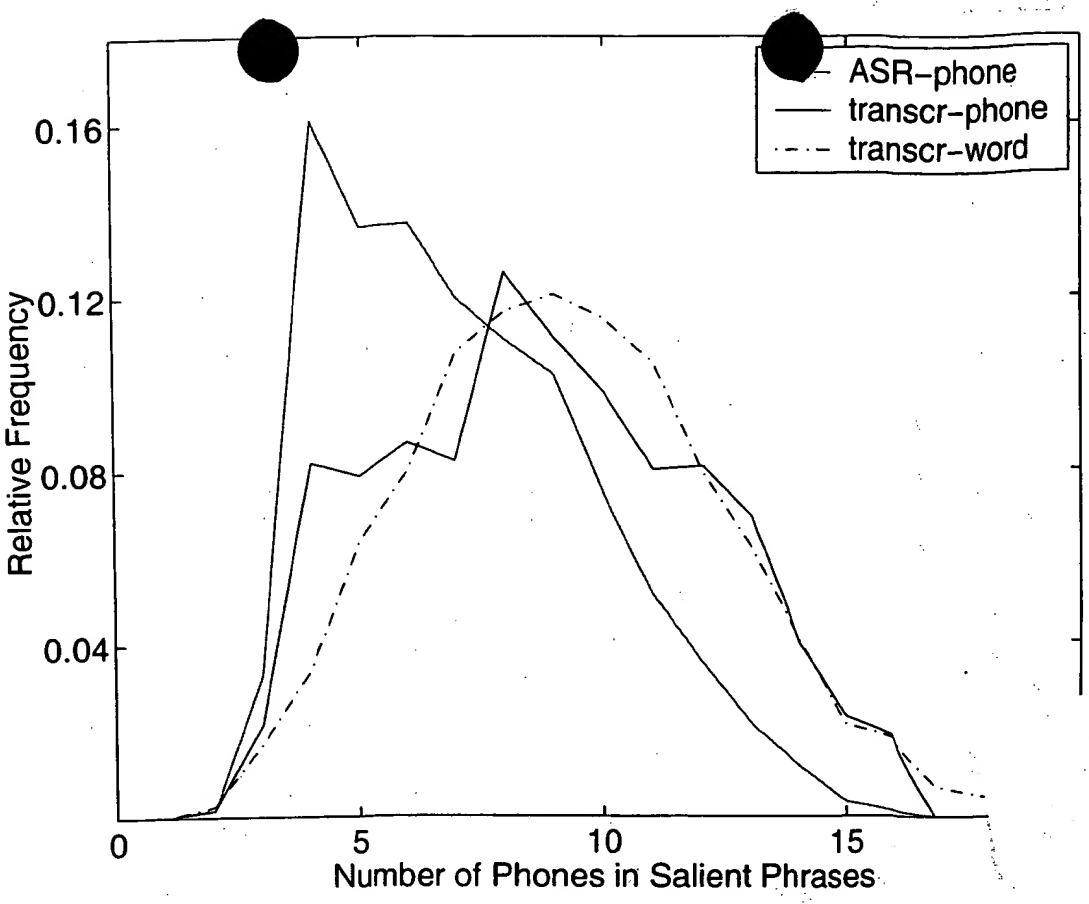
ASR–phone
transcr–phone
transcr–word
Relative Frequency
0.16
0.12
0.08
0.04
0
5
10
15
Number of Phones in Salient Phrases

Fig. 12

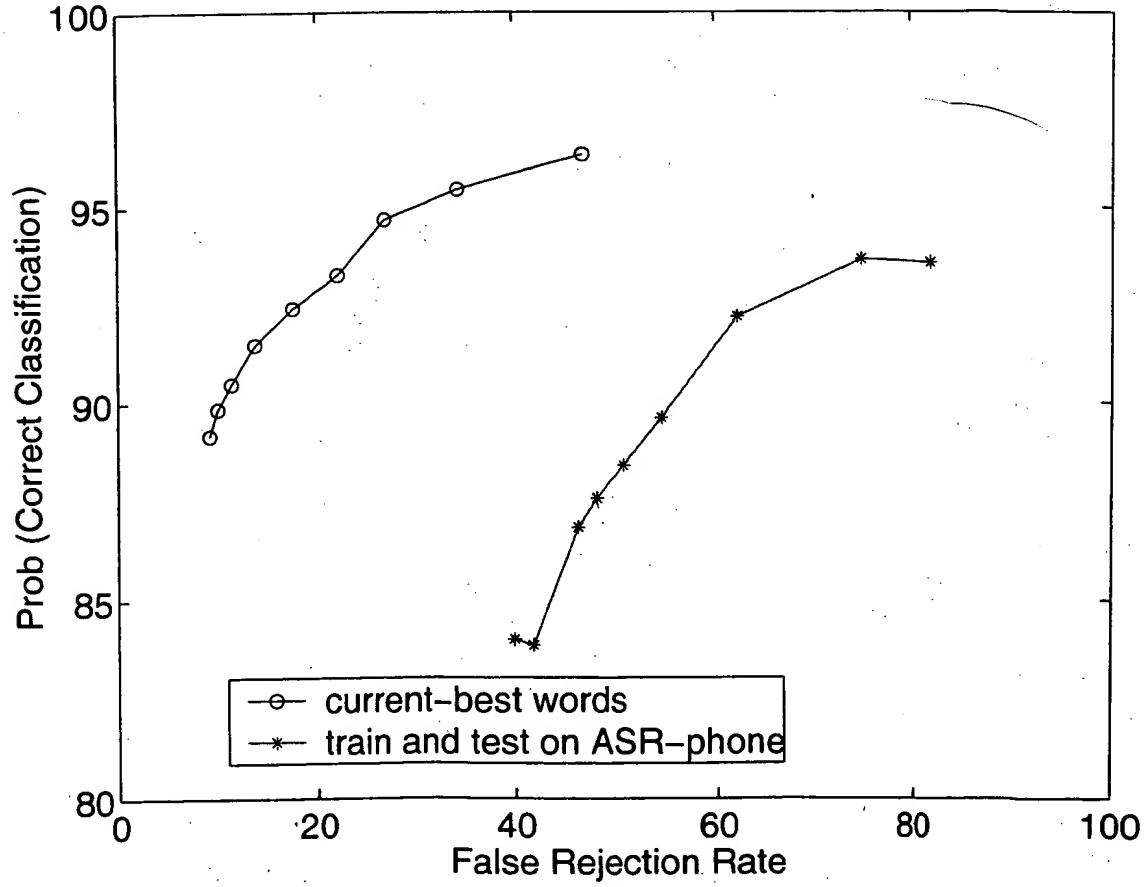
Prob (Correct Classification)
100
95
90
85
80
0
20
40
60
80
100
False Rejection Rate
current–best words
train and test on ASR–phone

Fig. 13